误差环境中参数识辨前测量信息的熵描述 *

李 静 [1]，丁海洋 [1]，任学尧 [2]

(1. 国防科技大学 信息通信学院, 西安 710106; 2. 空军工程大学, 西安 710043)

摘 要：在逆问题的参数识辨中，测量信息包含的信息量直接影响参数的重构精度，如何度量测量信息中的信息量，对选取参数识辨所需的测量点个数具有重要的指导作用。根据贝叶斯方法，将体现先验信息的先验概率和反映测量信息的似然概率融合得到待识辨参数样本出现的概率，结合信息熵和最大熵计算评价因子，将评价因子用于描述参数识辨前带有误差的测量信息所包含的信息量。计算实例表明，该计算方法可以有效地描述误差环境中的测量信息，对实际应用中测量信息的选取具有重要的指导意义。

关键词：参数识辨；信息熵；贝叶斯方法；误差；评价因子

中图分类号：TP391   doi: 10.3969/j.issn.1001-3695.2018.01.0037

Entropy description of measured information with error before parameter identification

Li Jing[1], Ding Haiyang[1], Ren Xueyao[2]

(1 *College of Information & Communication*, *National University of Defense Technology*, *Xi'an* 710106, *China*; 2. *Air Force Engineering University*, *Xi'an* 710143, *China*)

Abstract: In the parameter identification of inverse problems, it was well-known that the information quantity contained in measurement information influences the reconstruction precision of parameter directly. How to describe the information quantity in measurement information plays an important role in the selection of the number of measurement points. According to the Bayesian method, the probability of the parameter sample to be identified is calculated by combining the priori probability reflecting prior information and the likelihood probability reflecting measurement information. The evaluation factors are calculated by the maximum entropy and information entropy of the parameter. This paper introduced the information entropy to describe the information quantity of measurement information with error before parameter identification. Numerical tests showed that this computing method proposed could describe the measured information with error effectively and determine the measured information in practice.

Key words: parameter identification; information entropy; Bayesian method; error; evaluation factor

0 引言

参数识辨是逆问题中的一类重要问题，在逆问题的工程应用中，参数识辨前会提供大量的测量信息，如果这些数据准确，为了提高参数识辨的精确度，所用的测量信息越多越好；但是，在误差环境中，这些测量信息往往含有一定的测量误差，其数量的增多不仅会加大计算成本，更重要的是会导致重构计算中误差的积累。所以，在逆问题的计算过程中，当获得测量信息之后，通常需要先对大量数据进行筛选，即进行参数识辨前大量数据的分析和评估，以减小重构计算量的压力和重构误差。在参数识辨问题的研究中，已有的大量文献主要针对重构方法进行了讨论[1~3]，而忽略了重构前必须解决的重要问题——测量信息的选取。目前，针对这一问题本文作者已经提出了一种有效的数学方法——测量信息的熵描述[4]。

与过去文献中熵的用法[5,6]相比，测量信息的熵描述将熵用在参数识辨前测量信息的描述中[4]。本文针对参数识辨前大量带有误差的测量信息，将信息熵用于描述测量信息所包含的信息量，结合贝叶斯方法[7~9]，采用待识辨参数分布的后验概率定义参数样本出现的概率，计算出误差环境中不同数量的测量信息对应的评价因子，通过比较评价因子和重构误差，说明评价因子可以用来判断参数识辨前实际所需的测量信息，解决逆问题应用中的首要问题。不失一般性，本文在最后引入一实例对误差环境中测量信息的描述方法进行说明。计算结果表明：可以根据评价因子较好地进行逆问题中带误差的测量信息的描述。

收稿日期：2018-01-20；**修回日期：**2018-03-03 **基金项目：**国家自然科学基金青年基金资助项目（61301135）

作者简介：李静（1981-），女，山西临汾人，讲师，博士，主要研究方向为电磁兼容及电磁场中逆问题的研究（27176610@qq.com）；丁海洋（1980-），男，副教授，博士，主要研究方向为无线通信；任学尧（1983-），男，讲师，博士，主要研究方向为无线通信.

1　评价因子的定义以及测量信息的描述

1.1　最大熵的确定

在信息论中，信息熵就是随机事件不确定性大小的度量[10]。若逆问题的数学模型用算子方程表示为：$Af=y$，f 是待识辨参数，y 是测量项，它关于 f 的有限次测量值是 $y_1,y_2,\cdots,y_m$，A 是将 f 映射到测量值 $y_1,y_2,\cdots,y_m$ 的正向算子。如果随机事件 f 具有多种可能解 $f_1,f_2,\cdots,f_n$，每个解出现的概率为 $P_1,P_2,\cdots,P_n$，则该随机事件对应的信息熵为：

$$H(f)=-\sum_{i=1}^{n}P_i\ln P_i=\sum_{i=1}^{n}P_i\ln\frac{1}{P_i} \tag{1}$$

但是，如何量化测量信息对 P_i 的影响呢？根据贝叶斯方法，本文采用随机事件的后验概率来定义 P_i [11]。这样，就可以得到参数 f 在已知不同测量信息后的信息熵。

根据贝叶斯假设[7]，当无任何先验信息时，参数 f 可能的解 $f_1,f_2,\cdots,f_N$ 对应的先验概率 $\pi_i(f)$（$i=1,2,\cdots,N$）应在 f 的取值范围 D 内是"均匀"分布的，所以 $\pi_i(f)$ 可写成

$$\pi_i(f)=\begin{cases}\dfrac{1}{N} & 当 f\in D\\ 0 & 当 f\notin D\end{cases}\quad i=1,2,\cdots,N \tag{2}$$

其中:N 为参数 f 的样本数。本文令 $H_{\max}(f)$ 为 f 对应的最大熵(无任何先验信息时的熵)[4]，则 $H_{\max}(f)$ 为

$$H_{\max}(f)=\sum_{i=1}^{N}\pi_i(f)\ln\frac{1}{\pi_i(f)}=\sum_{i=1}^{N}\frac{1}{N}\ln\frac{1}{\frac{1}{N}}=\ln N \tag{3}$$

1.2　评价因子及测量信息的描述方法[4]

贝叶斯方法的基本思想是将先验概率和似然概率融合，得到问题的后验概率。根据贝叶斯方法的原理，有贝叶斯公式：

$$\pi_i(f|y_1,\cdots,y_m)=\frac{\pi_i(f)\pi_i(y_1,\cdots,y_m|f)}{\sum_{i=1}^{N}\pi_i(f)\pi_i(y_1,\cdots,y_m|f)} \tag{4}$$

式(4)右端的分母是一个常数，于是 $\pi_i(f|y_1,\cdots,y_m)$ 正比于分子：

$$\pi_i(f|y_1,\cdots,y_m)\propto\pi_i(f)\pi_i(y_1,\cdots,y_m|f) \tag{5}$$

其中：f 是待识辨参数；$\pi_i(f)$ 是 f 的先验概率，体现了其先验信息；$y_1,\cdots,y_m$ 是在误差环境中获得的测量信息；$\pi_i(y_1,\cdots,y_m|f)$ 是包含了测量信息的似然概率；$\pi_i(f|y_1,\cdots,y_m)$ 是 f 的后验概率，用它来表示式（1）中的 P_i，它综合了参数 f 的先验信息和 $y_1,\cdots,y_m$ 提供的测量信息，反映的是参数 f 的综合信息。该后验概率定义在整个解空间上，表示问题的"完全"解，也就是由所有可能出现的解组成的解集中任何一个解出现的概率。

将式（2）与（5）结合，式（5）可简化为

$$\pi_i(f|y_1,\cdots,y_m)\propto\pi_i(y_1,\cdots,y_m|f) \tag{6}$$

此时，参数 f 对应的信息熵为

$$H_m(f)=\sum_{i=1}^{N}[\pi_i(y_1,\cdots,y_m|f)\ln\frac{1}{\pi_i(y_1,\cdots,y_m|f)}] \tag{7}$$

为了描述似然概率的本质，令

$$\pi_i(y_1,\cdots,y_m|f)\propto\exp(-\tau_i)\ [4] \tag{8}$$

而误差 τ_i（>1 时）可由算子方程左右两端差值的范数给出[11]：

$$\tau_i=\frac{\|Af_i-y\|}{m-2} \tag{9}$$

综合式（8）与（9），代入式（7）有：

$$H_m(f)=\sum_{i=1}^{N}\exp(-\tau_i)\ln\frac{1}{\exp(-\tau_i)} \tag{10}$$

此式即为逆问题中已知 m 个测量数据后参数 f 信息熵的最终表达式。

通常对于实际问题中测量信息的选取，必须要对测量信息给出统一的数学方法进行描述[4]。定义：

$$\eta=\ln|\frac{H_{\max}(f)-H_m(f)}{H_{\max}(f)}| \tag{11}$$

为已知 m 个测量数据后的评价因子。它可以在待识辨参数的真实值未知的情况下，代替重构误差对参数的计算结果进行评价，并可以对参数识辨中实际所需的测量信息进行分析。

一般情况下，参数 f 的样本空间 N 大，不能对所有样本进行逐一计算，所以采用马尔可夫链蒙特卡罗法——Gibbs 抽样器对总样本抽样，产生一组服从后验概率 $\pi_i(f|y_1,\cdots,y_m)$ 分布的样本。下面用一个具体实例进行分析说明上述方法的有效性。

2　实例分析

利用一维的参数识辨模型[4]来说明利用上述方法进行测量信息的选取。该模型为

$$\begin{cases}\dfrac{d^2y(x)}{dx^2}+f(x)\cdot y(x)=0 & x\in[-1,1]\\ y(x)=1.71653 & x=\pm1\end{cases} \tag{12}$$

其中：$f(x)$ 为待识辨参数，$y(x)$ 为测量项。当 $f(x)$ 的解析式已知：

$$f(x)=\cos(x)-\sin^2(x)\qquad x\in[-1,1] \tag{13}$$

时，结合式（12）可以得到 $y(x)$ 的真实分布，如图 1 所示。

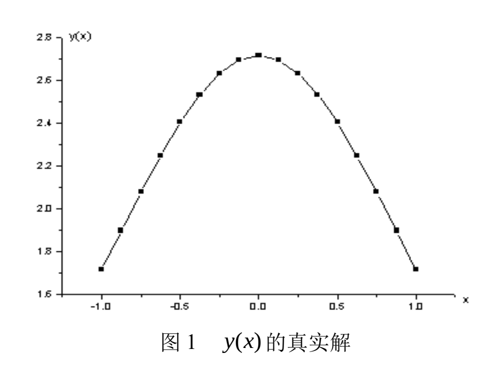


图 1　$y(x)$ 的真实解

本文将 $y(x)$ 沿 x 轴在[-1,1]内按 1/8 间隔均匀离散，共得到

17 个测量数据，给这些测量信息分别加入±3%、±5%、±10%和±15%的相对误差后，将其作为不同误差环境中的测量信息 $y^*(x)$。

为了验证前面关于逆问题参数识辨前测量信息的描述方法，下面本文假设 $f(x)$ 的解析式未知。已知 $y(x)$ 的测量信息 $y^*(x)$，将用有限差分法求得的 $f(x)$ 的解 $f^c(x)$ 与其解析解比较，有参数 $f(x)$ 的重构误差：

$$\delta=|\frac{f(x)-f^c(x)}{f(x)}|\times 100\% \tag{14}$$

再将重构误差与评价因子进行对比，从而说明用评价因子能够代替重构误差较好地描述不同误差环境中重构所需的测量信息。评价因子越大，重构误差越高，说明参数的重构精度越低，则重构所用的测量信息不足或不够准确。

与 $y(x)$ 相同，将 $f(x)$ 沿 x 轴在[-1,1]内均匀离散，离散间隔 $\Delta x=1/8$，则离散分量的个数 $k=17$。根据先验信息，$f(x)$ 的取值范围 $D=[-1/5,1]$，再将 $f(x)$ 在 $[-1/5,1]$ 内按 $\Delta f=2/15$ 的间隔等间隔离散，则离散分量的个数 $l=10$。则两次离散后 $f(x)$ 的总样本数 $N=l^k=10^{17}$。

本例中，随机抽取样本数 $M=2\times10^5$，根据式（3），$f(x)$ 对应的最大熵 $H_{\max}(f)=\ln M=\ln 2\times10^5=12.2061$。然后运用有限差分对 $\frac{d^2y(x)}{dx^2}$ 展开，再将式（12）代入式（9），此时 τ_i 即为

$$\tau_i=\frac{\|[y(x_j+2\Delta x)-2y(x_j)+y(x_j-2\Delta x)]/(2\Delta x)^2+f_i(x_j)\cdot y(x_j)\|}{m-2} \quad i=1,\cdots,M;\quad j=1,\cdots,m \tag{15}$$

其中，m 是测量数据的个数。

对于测量信息 $y^*(x)$，分别将 $f(x)$ 的 M 个样本代入式（15）求得 M 个 τ 值，将 τ_i 代入式（10）（11），便可得到 m 个测量数据对应的评价因子，结果如图 2 和 3 所示。由于已知测量信息是离散的，图 2 和 3 中的重构误差是在 $y(x)$ 相应测量点上 $f(x)$ 重构误差的平均。

图 2 和 3 中测量点的个数是按照图 1 中 $y^*(x)$ 的离散分量从左至右依次递增。通过图 2 可以看出，在没有测量误差的情况下，评价因子能够较好地反映重构误差的变化规律，并且测量点数越多，评价因子和重构误差都越小。

观察图 3 的（a）（b）不难看出，当加入误差之后，已知的测量信息越多，其对应的重构误差和评价因子并不一定越小（17 个测量点时的重构误差和评价因子并不是最小）。这主要是因为，利用带误差的测量信息进行重构时，在测量信息增多的同时，其携带的误差有可能导致重构误差的积累。同时图 3 还显示，重构误差和评价因子均会随着相对误差的增加而增大，但是，当相对误差不断变大时，图（b）中重构误差的变化越来越剧烈，而图（a）中评价因子的变化却是越来越平缓。由此可以得出这样的结论：当相对误差过大时，评价因子无法很好地反映重构误差地变化规律，不能用来进行测量信息的描述。通过图 3 可以得到这样的结论：测量信息的个数为 5 时，是评价因子最小和重构误差最低的时候，因此，该重构需要的测量信息的个数是 5。与 17 个甚至更多的测量点相比，利用 5 个测量点进行重构可以节约一定的计算成本。

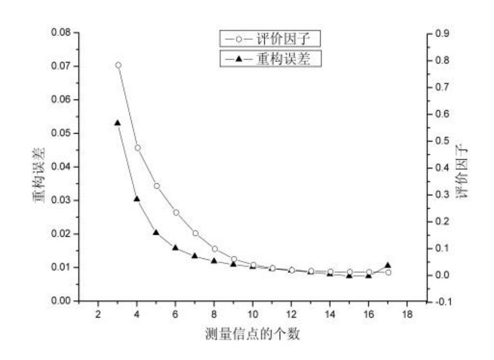
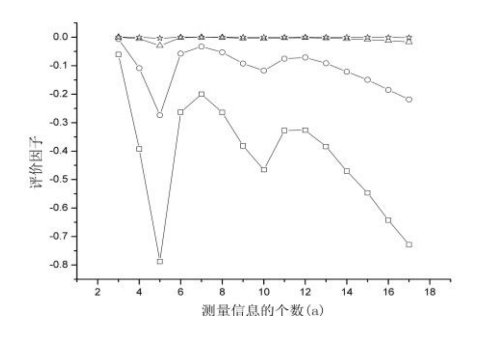


图 2　无测量误差时,不同数量的测量信息对应的评价因子和重构误差



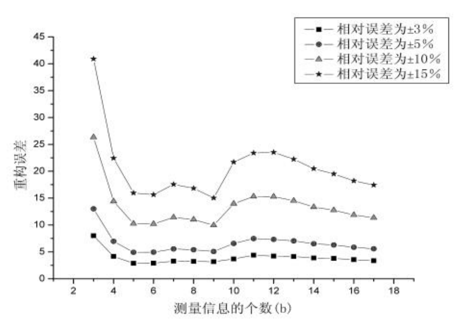


图 3　不同的误差环境中，不同数量的测量信息对应的评价因子和重构误差

因此，在逆问题的工程应用中，参数识辨前利用评价因子对误差环境中的测量信息进行描述，既可以较准确地完成参数识辨，又可以节省大量的计算成本。

3　结束语

在逆问题的工程应用中，测量信息的选取是一个必须解决的首要问题。本文针对实际应用中的这一问题，首次提出如何用评价因子代替重构误差对误差环境中参数识辨前的测量信息进行描述。同时，将信息熵引入到逆问题中，应用贝叶斯理论，将待识辨参数的先验信息描述为先验概率，测量信息表示为似然概率，二者共同得到其后验概率。再结合熵的计算公式，得到已知不同数量的测量信息时参数的信息熵。最后利用待识辨参数的最大熵和信息熵定义出不同数量的测量信息对应的评价因子。在本文的最后，通过具体的实例分析表明：相对误差不是很大时，用信息熵表示的评价因子能够较好地反映重构误差

的变化规律，可以用来完成逆问题中测量信息的描述。此方法不仅能满足重构误差低的要求，还可以大大降低计算量。该项研究对逆问题的工程应用具有重要的指导意义。

参考文献：

[1] Khulbe M, Tripathy M R, Parthasarathy H. Wavelet-based method for nonlinear inverse scattering problem using least mean square estimation [C]// Proc of Progress in Electromagnetics Research Symposium. [S. l.] : Spring, 2017: 496-501.

[2] Yalcin E, Ozdemir O, Taskin U. Sparsity based regularization for microwave imaging with NESTA algorithm [C]// Proc of IEEE Conference on Antenna Measurements & Applications. 2017: 282-283.

[3] Yang J, Hu D X, Wu Z R. Bayesian uncertainty inverse analysis method based on POME [J]. Journal of Zhejiang University: Engineering Science, 2006, 40 (5): 810-815, 835.

[4] 李静，黄卡玛，赵翔. 数学物理逆问题中测量信息的熵描述 [J]. 中国科学 E 辑：技术科学, 2008, 38 (3): 469-475.

[5] Benoit P, Zubelli J P. On the inverse problem for a size-structured population model [J]. Inverse Problems, 2007, 23 (3): 1037-1052.

[6] Kim S K, Junq B S, Lee W II. An inverse estimation of surface temperature using the maximum entropy method [J]. International Communications in Heat and Mass Transfer, 2007, 34 (1): 37-44.

[7] 蔡洪，张士峰，张金槐. Bayes 实验分析与评估 [M]. 长沙：国防科技大学出版社, 2004: 1-15.

[8] 张尧庭，陈汉峰. 贝叶斯统计推断 [M]. 北京：科学出版社, 1991: 17-27.

[9] Fitzpatrick B G. Bayesian analysis in inverse problems [J]. Inverse Problems, 1991, 7 (5): 675.

[10] 王彬. 熵与信息 [M]. 西安：西北工业大学出版社, 1994: 121-150.

[11] Mohammad-Djafari A. Joint estimation of parameters and hyperparameters in a Bayesian approach of solving inverse problems [C]// Proc of International Conference on Image Processing. 1996: 473-476.